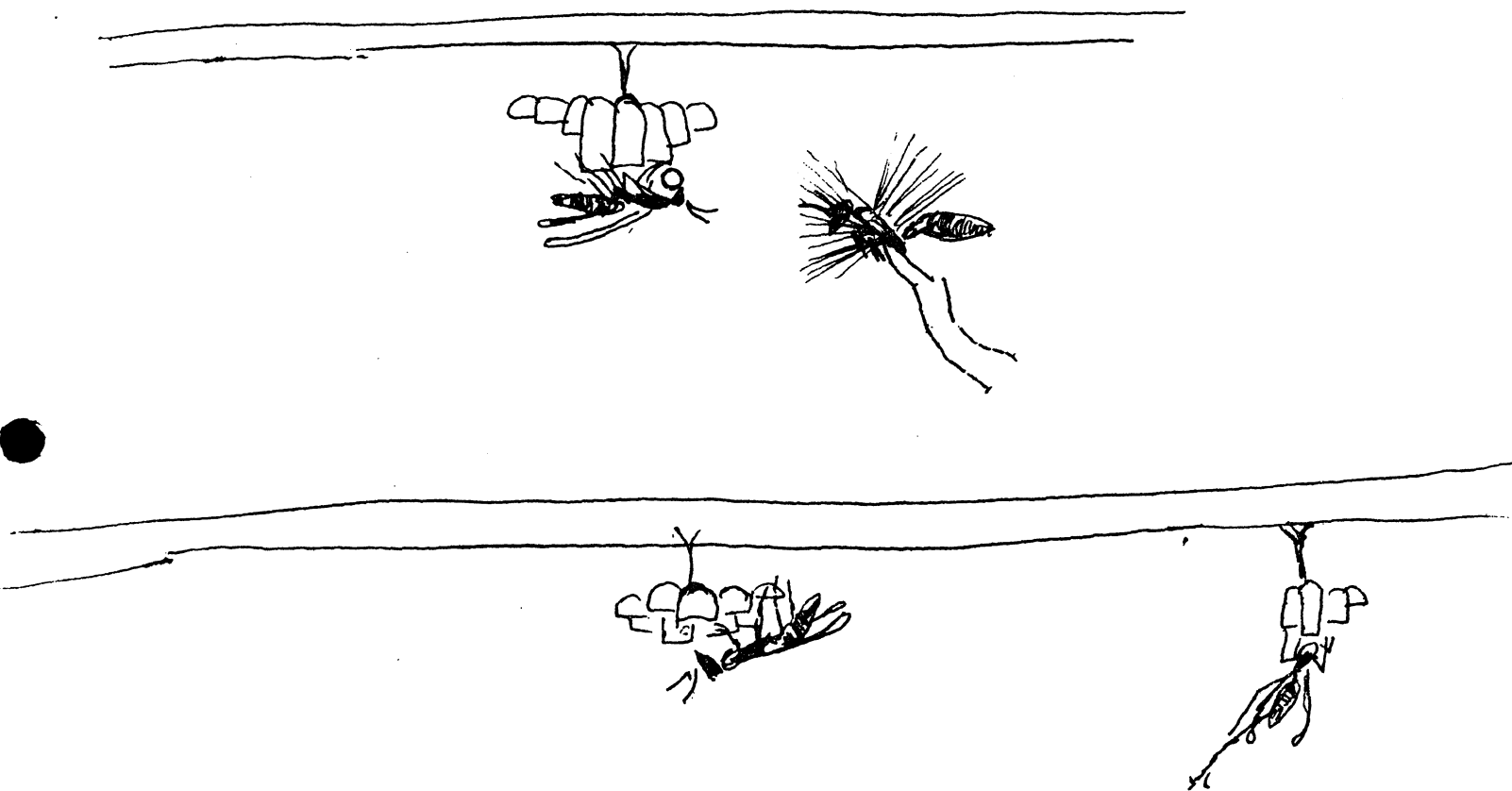


BU-344-M

A MODEL OF  
KIN SELECTION  
NOT INVOKING  
COEFFICIENTS OF  
RELATIONSHIP

M. J. ORLOVE



PAPER WASPS Polistes sp.  
SHOWING SOLITARY( BOTTOM)  
AND SOCIAL(TOP) COLONY FOUNDATION.

# ERRATA

page No. and location	error as it occurred	corrected error
-----------------------	----------------------	-----------------

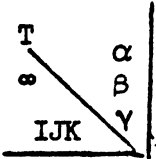
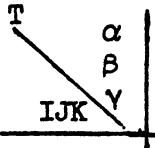
page 4 second line from top of page	then $P =$	then $\rho =$
--	------------	---------------

page 5 top line	$F_{BE}$	$F_{BE}$
--------------------	----------	----------

page 9 tenth line from top of page	$\delta$ (delta)	8(eight)
--	------------------	----------

page 14 table xvi	$A=1, V=0$	$A=0, V=1$
----------------------	------------	------------

page 15 table xvii	$A=1, V=0$	$A=0, V=1$
-----------------------	------------	------------

page 18 fig.4		
------------------	---	---

(extraneous infinity sign)

page 19 4th line from bottom of page	probability 0 sperm	probability a 0 sperm
--	---------------------	-----------------------

page 22	the legend" FACT #3" should not occur at the top of the page but right after eq. [ 11] .	
---------	---	--

page 26 Crozier	An. Nat.	Am. Nat.
--------------------	----------	----------

Hamilton	The geometrical evolution of social behavior	The Genetical Evolution of Social Behaviour
----------	--	---

## ACKNOWLEDGEMENTS

The author is indebted to the following persons.

Prof. Peter Brussard who read and criticized the first draft.\*  
Dr. John Emlen who read and discussed the first draft.\*\*\*\*\*  
Drs. Bruce Wallace and Ross MacIntyre for the same reasons.\*\*\*\*\*  
Dr. W. T. Federer and Mrs. VanOrder who made the seminar possible.\*\*\*\*\*  
Mrs. Seamon and Mrs. White for their typing of the manuscript.\*\*\*\*\*  
Dr. Vernon Wright who proofed the typed copy of the manuscript.\*\*\*\*\*  
Dr. L. C. Cole and Melvin Kreithen for their encouragement and who provided me with computer time.\*  
Dr. Henry D. Block for the same reasons.\*\*  
Dr. Frederick E. Warburton who sponsored the project in which version I was employed and whose advice was most useful.\*\*\*  
Dr. Howard Howland who provided the use of a small computer and encouragement.\*\*\*\*\*  
Dr. William L. Brown, Jr. who critically reviewed some ideas employed here and who aided in the preparation of the manuscript.\*\*\*\*\*  
Dr. Simon A. Levin who critically examined the model presented here, and who simplified some of the more complex algebraic expressions.\*\*  
Mr. John Colket who pointed out that some of my algebraic expressions could be more meaningful to others if they were expressed as Boolean expressions.\*\*\*\*\*  
Dr. Mary Jane West whose inspiration and encouragement made this paper possible.\*\*\*\*\*

---

\* Dept. of Ecology and Systematics, Cornell University.  
\*\* Dept. of Mathematics, Cornell University.  
\*\*\* Dept. of Biology, Barnard College, Columbia University.  
\*\*\*\* Dept. of Neurobiology and Behavior, Cornell University.  
\*\*\*\*\* Dept. of Entomology, Cornell University.  
\*\*\*\*\* Class of 1972, College of Engineering, Cornell University.  
\*\*\*\*\* Museum of Comparative Zoology, Harvard University.  
\*\*\*\*\* Dept. of Behaviour, Stoneybrook.  
\*\*\*\*\* Dept. of Genetics, Cornell University  
\*\*\*\*\* Dept. of Plant Breeding and Biometry, Cornell University.

### ABSTRACT

A model is presented in which a population is divided into colonies. The frequencies of the different kinds (genotypes) of colonies are represented in a state vector. The subscripts of the entries of this vector being descriptive of the genotype of the colonies. The general entry in the transition matrix employs Boolean functions of the digits of the base 2 representation of the subscripts. By this model kin selection is simulated without invoking coefficients of relationship, thus allowing us to estimate the coefficient of relationship during strong selection, at intermediate gene frequencies, and when the population is not in Hardy-Weinberg equilibrium.

### INTRODUCTION

The term kin selection was coined by Maynard Smith (1964). It denotes selection on genes through the effects these genes cause in individuals to have on the fitness of their relatives. Thus an animal substantially increasing the fitness of its relatives will increase the frequency of the genes causing this "altruistic" behaviour. H. G. Wells, et al (1931) were aware of this kind of selection. They explained how selection could operate on genes, only expressed in the phenotypes of sterile worker bees, by likening workers to the soma of an animal, and likening queens and males to the germ plasm. Although certain genes are not expressed in the phenotype of the germ plasm (queens, males) the fitness of these is influenced by the expression of these genes in the soma (workers). Hamilton (1964a) devised a complicated treatment of the subject which is beyond the scope of this paper. His conclusions are important however.

### HAMILTON'S RESULTS

An altruist is defined as an individual which increases the fitness of other members of its species. An altruist is on a limited resource budget. There is an opportunity cost of the fitness of the altruists neighbors in terms of the altruists fitness, i.e., a gain in fitness to a beneficiary is (often) accompanied by a loss in fitness on the part of his benefactor. This relationship is represented by the scalar K where

$$K = \frac{\text{Gain in beneficiary's fitness}}{\text{Loss in altruist's fitness}}$$

[1]

Assuming equal survivorship for altruist and beneficiary, we can rephrase

$$K = \frac{\text{Consequent increase in number of offspring produced by beneficiary}}{\text{Number of offspring given up by altruist}} \quad [2]$$

Altruism will be selected for if  $K > \frac{1}{r}$  where  $r$  is the fraction of genes identical by descent between the altruist and his beneficiary.  $r$  is called the coefficient of relationship. Various procedures exist for calculating  $r$  Hamilton (ibid) Falconer (1961), Malicott (in Falconer and Maynard Smith), Kempthorne (1957), Maynard Smith (1968), Crozier (1970) and Wright (1959).

In most cases the results are the same,  $r$  depends on the relative positions of the two animals in a pedigree.  $r$  is independent of gene frequency. Falconer (ibid) states that the value of  $r$  depends on how many generations back in time one wishes to calculate the pedigree.

When a gene for altruism is rare, selection slow, and the population is in equilibrium (little discrepancy in the gene frequencies, in the two sexes, for sex-linked gene), the procedures for calculating  $r$  yield results such that whether  $K > \frac{1}{r}$  determines whether selection is for altruism or against altruism.

At intermediate gene frequencies during rapid selection and changes in selective pressure  $K$  or  $r$  or both must be redefined before  $K > \frac{1}{r}$  determines the direction of selection. In order to avoid disagreements, I will use Greek letters  $\kappa$  and  $\rho$  in place of  $K$  and  $r$  respectively, using the right member of equation [2] to define  $\kappa$ . A model of kin selection can be employed which does not invoke  $\rho$ . If we assume

$$\kappa = \frac{\text{Consequent increase in number of offspring produced by beneficiary}}{\text{Number of offspring given up by altruist}}$$

regardless of gene frequency, or departure from equilibrium. Then  $\rho$  is set equal to whatever numerical value satisfies the condition  $\kappa = \frac{1}{\rho}$  when the (effective) gene frequency remains invariant, thus  $\rho$  varies with the state of the population and  $\kappa$  does not.

WEST'S USE OF  $\rho$

According to West's interpretation of Hamilton, West (1967) an individual's fitness [inclusive fitness Hamilton (ibid)] increases if the number of genes identical by descent to genes of that individual is increased. In general, an individual's fitness is increased if the frequency of genes identical [by descent or otherwise] is increased.

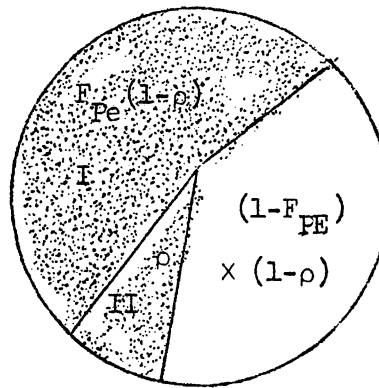
MY INTERPRETATION OF WEST

Thus, if we denote a potential altruist as EGO and his potential beneficiary as B, then  $P$  = the fraction of genes identical by descent between B and EGO; identity by descent being defined such that replication of B's genotype increases the frequency of genes identical (by descent or otherwise) to EGO's genes if and only if the number of genes identical by descent to EGO's genes is increased.

Let  $F_{PE}$  = the frequency of genes, identical (by descent or otherwise) to EGO's genes in the population.

Let  $F_{BE}$  = the fraction of genes identical (by descent or otherwise) between EGO and B.

Then B's genotype may be represented as a pie graph, viz.



The shaded portion  
=  $F_{BE}$

Figure 1.

Thus, segment II has a ratio to the rest of the graph of  $\frac{\rho}{1-\rho}$  and segment I has a ratio to segment III of  $\frac{F_{PE}}{1-F_{PE}}$ . Thus segments I and III represent a random sample of the population with respect to the occurrence of EGO's genes.

$$F_{BE} = F_{PE} \cdot (1-\rho) + \rho \quad [4]$$

and

$$\rho = \frac{F_{BE} - F_{PE}}{1 - F_{PE}} \quad [5]$$



We will save the calculation of  $F_{Pe}$  and  $F_{BE}$  for the next paper, and let it suffice that equation [4] justifies  $\rho$  as a function of the state of the population and  $\kappa$  as a supplied constant.

#### WEST'S WASPS

West (IBID) is interested in paper wasps, Polistes, which found colonies either as groups of sibling females or singly. When a group founds a colony a pecking order arises in which the top female lays the eggs and the others (auxilliaris) serve as workers. West's field observations (West (IBID) [page 1585, lower left]) on groups containing two wasps vs. "groups" containing one wasp indicate a value of  $\kappa$  sufficiently high for selection for altruism if  $r = .75$  (the conventional  $r$  between siblings in aculeate hymenoptera such as Polistes). That is, West's  $\kappa > \frac{1}{.75}$ . The question remains unanswered whether West's  $\kappa > \frac{1}{\text{West's } \rho}$ . This is not meant to criticize West, but rather to show the importance of  $\rho$  and the ability to estimate it from field data along with  $\kappa$ . It is not beyond the scope of the following model to do so with the data in West's paper, although this has not been done yet.

#### A MODEL OF KIN SELECTION

The model is limited to interactions between sibling hymenopterans ( $\rho$  is as if all genes were sex linked like West's Polistes wasps). Females are polyphenic, i.e., each female is of less fecund phenotype (worker, symbolized  $\text{♀}$ ) or of more fecund phenotype (queen symbolized  $\text{♀}$ ). This model considers selection on a single locus with two alleles. The phenotype with regards to  $\text{♀}$  and  $\text{♀}$  is either environmentally induced at random or controlled by a locus not linked to the one under consideration. As far as the locus under consideration is concerned, any genotype is equally frequent in the  $\text{♀}$  as it is in the  $\text{♀}$ , and a female zygote has the same probability of becoming a  $\text{♀}$  as it does of becoming a  $\text{♀}$  (50%) regardless of genotype. The gene under consideration is expressed in the  $\text{♀}$  phenotype only and controls altruism of the  $\text{♀}$  to a sibling  $\text{♀}$ .

GLOSSARY OF VARIABLE NAMES AND NOTATION IN THE MODEL

A - A variable equal to zero or one (i.e., a Boolean variable) which =1 when the ♀ inherits the gene from her mother from her maternal grandmother; =0 when the ♀ inherits the gene from her mother from her maternal grandfather.

AND symbolized  $\wedge$  - A function which operates on 0 and 1 to yield an answer 0 or 1. I.e.,  $\wedge$  has a domain of 0,1 and a range of 0,1, Iverson (1962), Berry (1969).  $\wedge$  is defined by the truth table.

Table I

X	Y	$X \wedge Y$
0	0	0
0	1	0
1	0	0
1	1	1

In conventional logic the range and domain of  $\wedge$  and other Boolean functions is denoted as T (true) and F (false) defined by truth table.

Table II

X	Y	$X \wedge Y$
F	F	F
F	T	F
T	F	F
T	T	T

We will abandon F and T for 0 and 1 respectively in this paper. The advantages will become evident as we proceed. As can be seen,  $\wedge$  is the equivalent of numeric multiplication  $\times$  as shown.

Table III

X	Y	$X \times Y$	$X \wedge Y$
0	0	0	0
0	1	0	0
1	0	0	0
1	1	1	1

B - A Boolean variable =0 when the ♀ inherits the gene from her mother from her maternal grandfather; =1 when the ♀ inherits the gene from her mother from her maternal grandmother.

BOOLEAN EXPRESSION - An expression = to 0 or 1. It may be generated by surrounding an equation with parentheses, viz.  $(A=B)$  ( $\Pi=3.14159$ ). If the equation is true the Boolean expression =1. If the equation is false the Boolean expression =0. This is another way of representing the Kronecker Delta function or the identity matrix I (defined as  $I_{i,j} = 1$  for  $i = j$ ;  $I_{i,j} = 0$  for  $i \neq j$ ). Figure 2 shows a  $4 \times 4$  identity matrix

$I_{i \backslash j}$	1	2	3	4
1	$I_{1,1}$ 1	$I_{1,2}$ 0	$I_{1,3}$ 0	$I_{1,4}$ 0
2	$I_{2,1}$ 0	$I_{2,2}$ 1	$I_{2,3}$ 0	$I_{2,4}$ 0
3	$I_{3,1}$ 0	$I_{3,2}$ 0	$I_{3,3}$ 1	$I_{3,4}$ 0
4	$I_{4,1}$ 0	$I_{4,2}$ 0	$I_{4,3}$ 0	$I_{4,4}$ 1

Figure 2.

or simply

1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

Figure 2. A  $4 \times 4$  identity matrix

Thus  $(P=Q)$  is equivalent to  $\delta_{P,Q}$  is equivalent to  $I_{P,Q}$  and  $(3=2+1)=1$  and  $(1=9)=0$ . The concept may be generalized to inequalities, Iverson (IBID) and Berry (IBID). Thus  $(7 > 1) = 1$  and  $(10 > 99) = 0$ . Boolean expressions may be combined with other Boolean expressions to form still larger Boolean expressions with the use of "AND" (symbolized  $\wedge$ ) and "OR" (symbolized  $\vee$ ), Iverson (IBID) and Berry (IBID). Thus  $((A=B) \wedge (C=D)) = 1$  if and only if A equals B and C equals D  $((A=B) \wedge (C=D)) = (A=B) \times (C=D)$  and  $((A=B) \vee (C=D)) = 1$  if A=B or if C=D.  $((A=B) \vee (C=D)) = ((A=B) - (C=D))^2 + (A=B) \times (C=D)$ . Look up "NOT" in this glossary for a further discussion of Boolean expressions.

OR symbolized  $\vee$  - A function with domain 0,1 and range 0,1 defined by the truth table, viz.

Table IV

X	Y	X $\vee$ Y
0	0	0
0	1	1
1	0	1
1	1	1

which is comparable to the truth table in terms of F and T

Table V

X	Y	X $\vee$ Y
F	F	F
F	T	T
T	F	T
T	T	T

$\vee$  is equivalent to the numeric product plus the square of the difference.

Table VI

X	Y	X $\vee$ Y	$(X-Y)^2 + XY$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	1

NOTE: Some prefer to simulate  $\vee$  with  $X+Y-XY$ . It will become clear later why  $(X-Y)^2 + XY$  is preferred here.

$N_{X,Y,Z}$  - A vector with Boolean subscripts which represents the population.

$X = 0,1$  ,  $Y = 0,1$  ,  $Z = 0,1$  .  $N$  could have been represented as a vector with one subscript; viz.  $N_M$ ,  $M = 0,1,2,3,4,5,6,7$ , and we could represent bits of  $M$ , i.e., digits of the base 2 representation of  $M$ . Where  $X$ ,  $Y$ , and  $Z$  are the bits of  $M$  and  $M = 4 \times X + 2 \times Y + Z$ . It is simpler to work with  $X$ ,  $Y$ , and  $Z$  and forget  $M$ , and when summing to write  $\sum_{X=0}^1 \sum_{Y=0}^1 \sum_{Z=0}^1$  in place of  $\sum_{M=0}^7$ . A colony is

made of a ♂ and a ♀ who are siblings. Since the locus is sex-linked one gene is always identical between them, i.e., there are three independent genes [Li (1955)] represented as  $X$ ,  $Y$  and  $Z$ . The alleles are represented as 0 and 1 giving rise to 8 evaluations of the 3 subscripts corresponding to the binary numbers from 0 to 7 which represent the 8 genotypes of the colonies. A colony is of genotype  $XYZ$  when the worker is of genotype  $X|Y$  and the queen is genotype  $Y|Z$ . This will be explained more clearly later.

NOT - Symbolized as overscore. Thus  $(\text{NOT } X) = \bar{X}$  and  $\text{NOT } (A=B) = (\overline{A=B})$ .  $\bar{X}=0$  when  $X=1$  and  $\bar{X}=1$  when  $X=0$ . Not has the following truth tables.

Table VII

$X$	$\bar{X}$
0	1
1	0

Table VIII

$X$	$\bar{X}$
F	T
T	F

and is equivalent to the numeric operation of subtraction (of a Boolean expression) from 1, viz.

$X$	$\bar{X}$	$1-X$
0	1	1
1	0	0

$\phi$  - The fraction of sperm carrying the altruistic gene.

QUEEN - Symbolized ♀. The phenotype which is aided by a ♂.

$Q_1$  - The number of offspring produced by an unaided ♀.

$Q_2$  - The extra offspring produced by a ♀ when she is aided by a ♂. An aided ♀ has  $Q_1 + Q_2$  offspring.

SD - A Boolean variable; in terms of simple dominance.

SD=0 - When altruism is dominant.

SD=1 - When selfishness is dominant. In this system V may be replaced by 1 and  $\wedge$  by 0. Thus  $A \vee B = A \vee B$  and  $A \circ B = A \wedge B$ . For example, De Morgan's laws can be stated as (op = 0,1),  $X \circ_p Y = \overline{\overline{X} \circ_p \overline{Y}}$ . Thus when selfishness is dominant SD = 1. When altruism is dominant SD = 0.  $V \equiv 1$  and  $\wedge \equiv 0$ .

w or W - The number of offspring produced by a selfish  $\varphi$  (a  $\varphi$  not aiding a  $\varphi$ ) an altruistic  $\varphi$  has zero offspring and aids a  $\varphi$ .

WORKER - A female symbolized  $\varphi$  in whose phenotype the gene for altruism is expressed, the less fecund phenotype.

### SOME MORE LOGIC

The operations  $\wedge$ , V and  $\neg$  were defined and their numeric equivalents given. Two more operations exist which should be familiar to everyone: equals, =, and not equals  $\neq$ . The latter is sometimes called exclusive or. Truth tables are shown in tables IX and X.

Table IX

X	Y	X=Y
0	0	1
0	1	0
1	0	0
1	1	1

Table X

X	Y	X $\neq$ Y
0	0	0
0	1	1
1	0	1
1	1	0

The numeric equivalent of  $\neq$  is the square of the difference.

Table XI

X	Y	$X \neq Y$	$(X-Y)^2$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	0	0

The numeric equivalent of  $=$  is one minus the square of the difference

Table XII

X	Y	$X=Y$	$1-(X-Y)^2$
0	0	1	1
0	1	0	0
1	0	0	0
1	1	1	1

DERIVING THE MODEL

The allele for altruism is denoted by zero, 0. The allele for selfishness is denoted by one, 1. Each colony has a ♂ and a ♀ so we can find the genotypes of the different kinds of colonies by combining the genotypes of a ♂ and a ♀ (each colony has one ♂ and one ♀ ).

Table XIII

♂	♀
0 0	0 0
0 0	0 1
0 0	1 1
0 1	0 0
0 1	0 1
1 1	1 1
1 1	0 0
1 1	0 1
1 1	1 1

Since the locus is sex-linked, two of these colony genotypes cannot exist, namely 0 0 1 1 and 1 1 0 0. Thus we can represent the colony genotypes as three columns of numbers since one gene is always identical between the ♂ and the ♀ . This gene will go in the middle column and represent the genotype of the male [hymenopteran males are haploid] who is the father of the ♂ and ♀ . The left column represents the contribution the ♂ gets from the mother and the right column represents the contribution the ♀ got from the mother. (See Table XIV.)



Table XIV

I = left column

J = middle column

K = right column

♂		♀
Ovum to become ♂	Sperm	Ovum to become ♀
I	J	K
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1
♂ genotype		♀ genotype

The above diagram shows the 8 kinds of colonies which can exist. The fact that there are 8 colonies shows that there are three bits [binary digits] of information in specifying the genotype of a colony. There is one locus, 3 genomes, and 2 alleles or  $1 \times 2^3$  kinds of colonies. This is what Li means by 3 independent genes existing in a system of this sort. Li (IBID, page 18).

There is a correspondence between the 4 column and 3 column system; see Table XV.

Table XV

	♂	♀	
	0 0	0 0	←
	0 0	0 1	←
NULL	0 0	1 1	
	0 1	0 0	←
* 0 1	0 1		←
0 1	1 1		←
NULL	1 1	0 0	
	1 1	0 1	←
	1 1	1 1	←

	♂	♀
	0 0 0	
	0 0 1	
	0 1 0 *	
	0 1 1	
	1 0 0	
	1 0 1 *	
	1 1 0	
	1 1 1	

4 columns                      3 columns

$$* \# 0 1 0 + \# 1 0 1 = \# 0 1 0 1;$$

0 1 0 and 1 0 1 produce the same offspring  
but they spring from different ancestors.

An altruistic ♂ has zero offspring and a selfish ♂ has  $w$  offspring. An altruistic ♂'s beneficiary has  $Q_1 + Q_2$  offspring and a selfish ♂'s "beneficiary" has  $Q_1$  offspring.

Assume selfishness dominant. Then the following diagram indicates how many offspring are produced by the ♂ and ♀.

Table XVI

$$SD = 1, A = 1, V = 0$$

Genotypes

Numbers of Offspring

♂	♀		♂	♀
I	J	K		
0	0	0	0	$Q_1 + Q_2$
0	0	1	0	$Q_1 + Q_2$
0	1	0	$w$	$Q_1$
0	1	1	$w$	$Q_1$
1	0	0	$w$	$Q_1$
1	0	1	$w$	$Q_1$
1	1	0	$w$	$Q_1$
1	1	1	$w$	$Q_1$

A similar diagram exists for altruism dominant and selfishness recessive.

Table XVII

$$SD = 0, \Lambda = 1, V = 0$$

Genotypes			Numbers of Offspring	
$\begin{array}{c} \text{♂} \quad \text{♀} \\ \text{I} \quad \text{J} \quad \text{K} \end{array}$			♂	♀
0	0	0	0	$Q_1 + Q_2$
0	0	1	0	$Q_1 + Q_2$
0	1	0	0	$Q_1 + Q_2$
0	1	1	0	$Q_1 + Q_2$
1	0	0	0	$Q_1 + Q_2$
1	0	1	0	$Q_1 + Q_2$
1	1	0	w	$Q_1$
1	1	1	w	$Q_1$

When selfishness is dominant either I or J must equal one before a ♂ of genotype IJ can have w offspring; otherwise she has 0 offspring.

Anything multiplied by 1 = itself. Anything multiplied by 0 = 0. Here is where truth values of 0 and 1 become useful. When selfishness is dominant a ♂ has  $w \times (I \vee J)$  offspring and a queen has  $Q_1 + Q_2 \times \overline{(I \vee J)}$  offspring; i.e., a worker has  $w \times (I \wedge J)$  offspring and a queen has  $Q_1 + Q_2 \times \overline{(I \wedge J)}$  offspring. When altruism is dominant a ♂ has  $w \times (I \wedge J)$  offspring and a ♀ has  $Q_1 + Q_2 \times \overline{(I \wedge J)}$  offspring; i.e., the ♂ has  $w \times (I \circ J)$  offspring and the ♀ has  $Q_1 + Q_2 \times \overline{(I \circ J)}$  offspring.

In general the ♂ has  $w \times (I \text{ sd } J)$  offspring and the ♀ has  $Q_1 + Q_2 \times \overline{(I \text{ sd } J)}$  offspring. If the argument is not clear at this point, look up the following items in the glossary: AND, OR, BOOLEAN EXPRESSION, NOT, and SD, in that order. Then read this section again. Then show that Tables XVI and XVII are in agreement with  $w \times (I \text{ sd } J)$  and  $Q_1 + Q_2 \times \overline{(I \text{ sd } J)}$  as numbers of offspring for the ♂ and ♀, respectively.

Now that we know how many offspring are produced by each member of the colony as a function of the ♂'s genotype, the question arises: What genotypes will these

offspring be and what will be the genotypes of the colonies they form? Let us examine the offspring of a ♀ of genotype IJ. The ova of this ♀ will be of genotype I and of genotype J. We can imagine the ova pairing to form colonies in which one becomes the ♂ and the other becomes the ♀. Which becomes ♂ is randomly determined. It is not important to know at this stage the genotype of the male whose sperm fertilizes these eggs. Just as random mating can be represented as the random pairing of gametes from a population, thus may the random pairing of sisters (with a common haploid father, all of whose sperm are identical) be represented as the random pairing of ova produced by a female. Since mating is random in this model, these ovum pairs will be fertilized by sperm of genotype 0 (zero) with frequency  $\phi$ , and by sperm with genotype 1 with frequency  $(1 - \phi)$ . At this point it might be noted that although many hymenoptera mate with several males in succession the sperm of one male is probably used up before the female uses the sperm from another. This is definitely true in the honeybee *Apis mellifera* L. whose queens store sperm for years but do not mix the wads from different males but exhaust them sequentially, Taber (1955). Thus two females, the same age, usually have the same father. A ♀ of genotype IJ produces ova of genotype I and of genotype J in equal frequency. These ova form pairs II, IJ, JI, JJ.

Table XVIII

♀ I J	Ovum Pairs			
	I I	I J	J I	J J
0 0	0 0	0 0	0 0	0 0
0 1	0 0	0 1	1 0	1 1
1 0	1 1	1 0	0 1	0 0
1 1	1 1	1 1	1 1	1 1

The following figure (Figure 3) shows gene flow from a ♀ to her daughter colonies. The ♀ is of genotype IJ and the daughter colonies are designated as  $\alpha \beta \gamma$  where  $\alpha$  corresponds to I,  $\beta$  corresponds to J, and  $\gamma$  to K, i.e., the genotypes of colonies in the next generation are referred to as  $\alpha \beta \gamma$ , not I J K. The following genotypes are possible among the offspring of a female of genotype IJ.

General Form  
 $\alpha = I \cdot A + J \cdot \bar{A}$   
 $\gamma = I \cdot B + J \cdot \bar{B}$

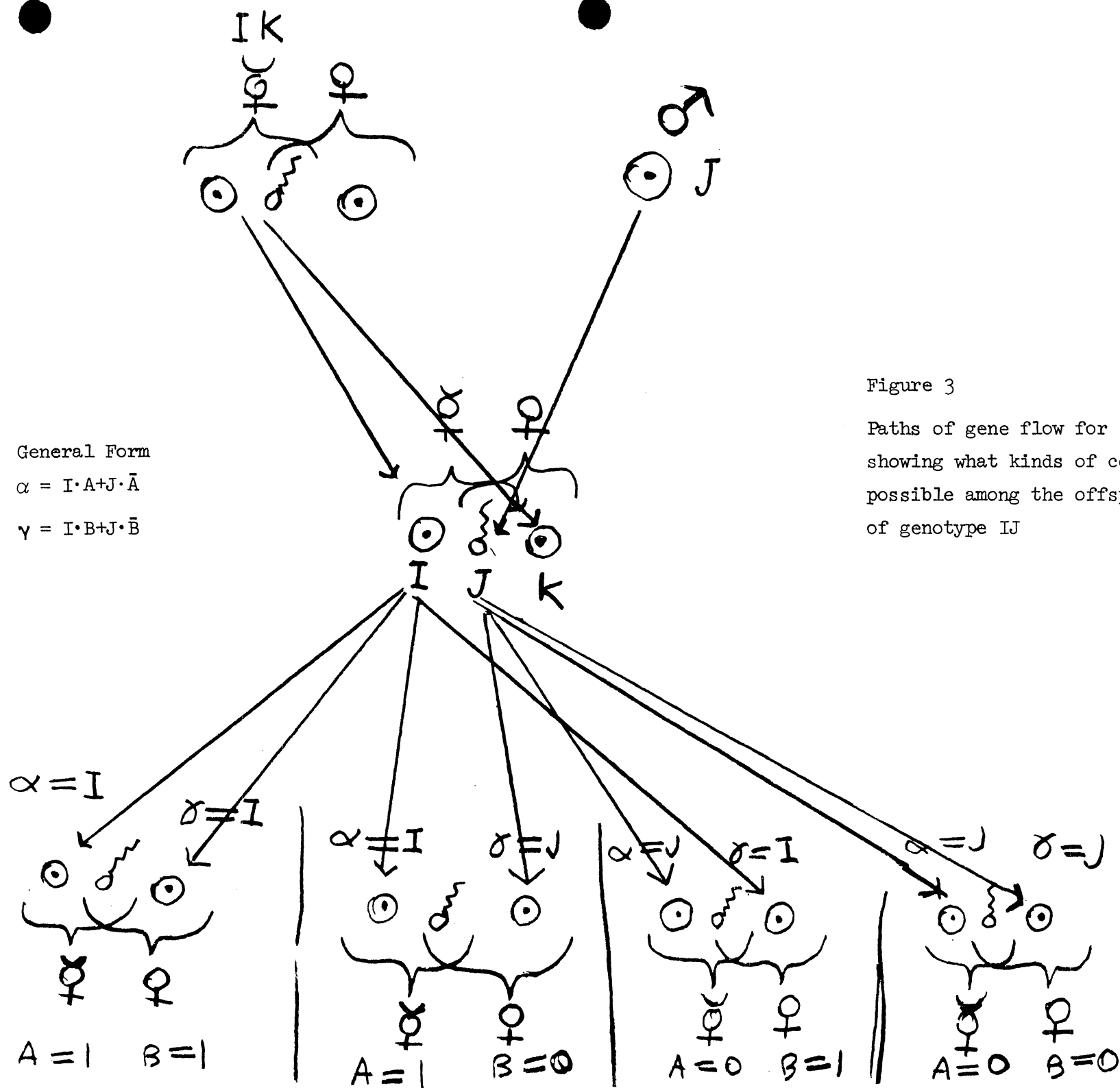


Figure 3

Paths of gene flow for 3 generations showing what kinds of colonies are possible among the offspring of a female of genotype IJ

SEE GLOSSARY TO LOOK A AND B UP

$\alpha = I$                       when  $A = 1$   
                                    and  $B = 1$

$\beta = 0,1$

$\gamma = I$

---

$\alpha = I$                       when  $A = 1$   
                                    and  $B = 0$

$\beta = 0,1$

$\gamma = J$

---

$\alpha = J$                       when  $A = 0$   
                                    and  $B = 1$

$\beta = 0,1$

$\gamma = I$

---

$\alpha = J$                       when  $A = 0$   
                                    and  $B = 0$

$\beta = 0,1$

$\gamma = I$

---

Temporarily we can let  $Q_1$  and  $Q_2 = 0$ , then we can leave them out of our equations without distorting the truth.

Let  $N_{I,J,K}$  = the number of colonies of genotype IJK. Then we hope to set up a matrix  $T$  such that  $N' = N \cdot T$  [6] where  $N'_{\alpha,\beta,\gamma}$  is the number of colonies of genotype  $\alpha,\beta,\gamma$  in the following generation.

This matrix appears in Figure 4; the reader should cover some of the entries and try to guess what they are.

T 8	I J K	$\alpha$							
		$\beta$	$\gamma$						
		0	1	0	1	0	1	0	1
		0	0	1	1	0	0	1	1
		0	0	0	0	1	1	1	1
0 0 0		0	0	0	0	0	0	0	0
0 0 1		0	0	0	0	0	0	0	0
0 1 0		$w \cdot SD \cdot \phi / 4$	$w \cdot SD \cdot \phi / 4$	$w \cdot SD \cdot (1-\phi) / 4$	$w \cdot SD \cdot (1-\phi) / 4$	$w \cdot SD \cdot \phi / 4$	$w \cdot SD \cdot \phi / 4$	$w \cdot SD \cdot (1-\phi) / 4$	$w \cdot SD \cdot (1-\phi) / 4$
0 1 1		↓	↓	↓	↓	↓	↓	↓	↓
1 0 0		↓	↓	↓	↓	↓	↓	↓	↓
1 0 1		↓	↓	↓	↓	↓	↓	↓	↓
1 1 0		0	0	0	0	0	$w \cdot \phi$	0	$w \cdot (1-\phi)$
1 1 1		0	0	0	0	0	$w \cdot \phi$	0	$w \cdot (1-\phi)$

Figure 4 a transition matrix.

$T_{I,J,K, \alpha, \beta, \gamma}$  is the number of  $\alpha, \beta, \gamma$  colonies made by an I, J, K colony in one generation.

$Q_1$  and  $Q_2$  temporarily = 0 so they do not figure in this matrix.

The top two rows of Figure 4 (rows 000 and 001) = 0 because (in addition to other factors equalling 0 in some of the entries)  $(I \neq J) = 0$  so  $\check{\check{q}}\check{\check{q}}$  (plural of  $\check{q}$ ) are having 0 offspring. In rows 010, 011, 100 and 101 the  $\check{q}$  is a heterozygote and a  $\check{q}$  has  $w$  offspring if selfishness is dominant, i.e.,  $SD = 1$  and 0 offspring if  $SD = 0$ ; thus the  $w \cdot SD$  in each entry in these rows.

In rows 110 and 111 the zero entries occur because the  $\check{q}$  of genotype 1 1 cannot produce offspring which form colonies;

$$\alpha = 0, \beta \gamma = 0$$

$$\alpha = 1, \beta \gamma = 0 \quad \beta = 0, 1$$

$$\alpha = 0, \beta \gamma = 1$$

Such a  $\check{q}$  can only produce

$$\alpha = 1, \beta, \gamma = 1.$$

In other words

$$T_{I,J,K, \alpha, \beta, \gamma} = 0$$

When the following sum of Boolean expressions = 0

$$\sum_{A=0}^1 \sum_{B=0}^1 \frac{((\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}))}{4}$$

This sum is a factor in the general formula for  $T_{I,J,K, \alpha, \beta, \gamma}$  when  $I \neq J$  the  $\check{q}$  makes all kinds of ovum pairs but only one iteration of the above sum = 1, thus the division by 4 in entries in rows 0 1 0, 0 1 1, 1 0 0, 1 0 1.

Some columns have a factor of  $\phi$ , these are those where  $\beta = 0$ , since the probability 0 sperm fertilizes an ovum pair = the frequency of altruistic sperm =  $\phi$ . When  $\beta = 1$  the column entries contain a factor of  $1 - \phi$ , since  $1 - \phi$  is the frequency of selfish sperm. So the factor involving  $\phi$  is always  $|\beta - \phi|$  and the general formula (when  $Q_1 = 0$  and  $Q_2 = 0$ ) for  $T_{I,J,K, \alpha, \beta, \gamma}$  is:



$$0 \leq w < \infty$$

$$Q_1 = 0$$

$$Q_2 = 0.$$

$$T_{I,J,K, \alpha, \beta, \gamma} = w \cdot (I \text{ s } J) \cdot |\beta - \phi| \cdot \sum_{A=0}^1 \sum_{B=0}^1 \frac{((\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}))}{4} \quad [7]$$

substituting [7] into eq. [6] gives

$$N'_{\alpha, \beta, \gamma} = \sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \sum_{B=0}^1 w \cdot (I \text{ s } J) \cdot N_{I,J,K} \cdot |\beta - \phi| \cdot ((\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}))/4 \quad [8]$$

the constant w is inside the summation sign for good reason.

Now let  $w = 0$  and let  $Q_1 \neq 0$  and  $Q_2 \neq 0$ . We can write equations in which the w terms can be left out and the  $Q_1$  and  $Q_2$  terms are written.

We could set up a transition matrix similar to Figure 4 with  $Q_1$  and  $Q_2$  terms and no w terms. These would be a factor of  $Q_1$  whenever  $(I \text{ s } J) = 0$  and a factor of  $Q_1 + Q_2$  whenever  $(I \text{ s } J) = 1$  the  $|\beta - \phi|$  factor would be the same.

Now the queens genotype is JK and  $\alpha$  and  $\gamma$  are compared to J and K not to I and J as before. Therefore

$$\sum_{A=0}^1 \sum_{B=0}^1 \frac{((\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}))}{4}$$

would be replaced by

$$\sum_{A=0}^1 \sum_{B=0}^1 \frac{((\alpha = K \cdot A + J \cdot \bar{A}) \wedge (\gamma = K \cdot B + J \cdot \bar{B}))}{4}$$

since we are interested in what  $\alpha, \beta, \gamma$  colonies come from a ♀ of genotype JK.

The general form of an entry in the matrix would be

$$w = 0, \quad 0 \leq Q_1 < \infty, \quad 0 \leq Q_2 < \infty$$

$$T_{I,J,K, \alpha, \beta, \gamma} = (Q_1 + Q_2 \cdot \overline{(I \text{ sib } J)}) \cdot |\beta - \phi| \cdot \sum_{A=0}^1 \sum_{B=0}^1 \frac{((\alpha = K \cdot A + J \cdot \bar{A}) \wedge (\gamma = K \cdot B + J \cdot \bar{B}))}{4} \quad [9]$$

substituting [9] into [6] gives us

$$N'_{\alpha, \beta, \gamma} = \sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \sum_{B=0}^1 (Q_1 + Q_2 \cdot \overline{(I \text{ sib } J)}) \cdot |\beta - \phi| \cdot N_{I,J,K} \cdot ((\alpha = K \cdot A + J \cdot \bar{A}) \wedge (\gamma = K \cdot B + J \cdot \bar{B}))/4 \quad [10]$$

If we sum equations [8] and [10] we get the formula desired for all values of  $w, Q_1$ , and  $Q_2$  but the notation gets messier. So we must alter EQ. [10] such that the sum will not look too different from [8] or [10]. We can take advantage of some facts:

1. In a sib pair either sibling could have been queen; queenness strikes without regard to genotype, i.e.,  $N_{I,J,K} = N_{K,J,I}$  always.

2. When summing over several variables, it makes no difference in what order we sum over which variable. i.e.,

$$\sum_A \sum_B = \sum_B \sum_A$$

3. I and K are bound variables. We can consistently replace all occurrences of either or both of them with some other variables. Thus we can substitute I for K and K for I simultaneously into [10]. This is the same as saying there can exist a paper identical to this one except wherever K occurs in this paper, I occurs in the other paper, and wherever I occurs in this paper, K occurs in the other paper.

The two papers would be equally true since they would be based on the same axioms, and the logic would be identical. Equation [10] from that paper would be excellent material for this paper and so substituting

$$N_{K,J,I} \text{ for } N_{I,J,K} \quad \text{FACT NO. 1}$$

and substituting

I for K

FACT NO. 2

and substituting

K for I

FACT NO. 3

Into EQ[10] we get

$$w = 0$$

$$0 \leq Q_1 < \infty$$

$$0 \leq Q_2 < \infty$$

[11]

$$N'_{\alpha, \beta, \gamma} = \sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \sum_{B=0}^1 (Q_1 + Q_2 \cdot \overline{(K \text{ and } J)}) \cdot N_{I, J, K} \cdot |\beta - \phi|$$

$$\cdot ((\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}))/4 .$$

Now we add Equation [8] to Equation [11], remembering to factor out the factor in  $N_{IJK}$  and  $|\beta - \phi|$ , etc. and to divide both sides by 2. We get

$$\begin{aligned} 0 &\leq w < \infty \\ 0 &\leq Q_1 < \infty \\ 0 &\leq Q_2 < \infty \end{aligned} \quad [12]$$

$$N'_{\alpha, \beta, \gamma} = \sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \sum_{B=0}^1 \left( w \cdot (I \text{ sd } J) + Q_1 + Q_2 \cdot \overline{(J \text{ sd } K)} \right) \cdot N_{I,J,K} \cdot |\beta - \phi|$$

$$\cdot \left( (\alpha = I \cdot A + J \cdot \bar{A}) \wedge (\gamma = I \cdot B + J \cdot \bar{B}) \right) / 8.$$

Since  $\phi$  changes every generation,  $T$  changes every generation, so this is not a true Markov chain. The vector  $N$  represents numbers rather than frequencies, but this can be remedied by dividing the vector by its sum.  $\phi'$  is calculated by dividing the number of zero ova produced by the total number of ova (♂♂ come from unfertilized ova). This is shown by the formula:

$$\phi' = \frac{\sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \left( w \cdot (I \text{ sd } J) + Q_1 + Q_2 \cdot \overline{(J \text{ sd } K)} \right) \cdot N_{I,J,K} \cdot \overline{(I \cdot A + J \cdot \bar{A})}}{\sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \left( w \cdot (I \text{ sd } J) + Q_1 + Q_2 \cdot \overline{(J \text{ sd } K)} \right) \cdot N_{I,J,K}}$$

$$\begin{aligned} 0 &\leq w < \infty \\ 0 &\leq Q_1 < \infty \\ 0 &\leq Q_2 < \infty \end{aligned} \quad [13]$$

where

This model which does not invoke  $\rho$  can be used to find  $\rho_0$  given  $N$  and  $\phi$ .

The effective gene frequency for a sex-linked gene =  $\frac{2\phi' + \phi}{3}$ , Wright (1959).  
Thus  $\frac{2\phi' + \phi}{3}$  should be invariant when  $\kappa = \frac{1}{\rho}$ . Thus

$$\frac{2\phi'' + \phi'}{3} = \frac{2\phi' + \phi}{3} \quad [14]$$

yields

$$2\phi'' = \phi' + \phi. \quad [15]$$

When  $\kappa = \frac{1}{\rho}$ ,  $\phi'$  need not =  $\phi$  as  $\phi^{\text{general}}$  goes through damped oscillations around effective  $\phi = \frac{2\phi' + \phi}{3}$  when the population is out of equilibrium.  $\phi''$ ,  $\phi'$ , and  $\phi$  are all functions of  $N$  and  $\phi$ .  $\phi = \phi$ ,  $\phi'$  is a function of  $N$ , and  $\phi''$  is a function of  $N'$  which is a function of  $N$  and  $\phi$ . The appropriate substitution would be too unwieldy to represent here.

By setting  $Q_2 = 1$ ,  $\left(\frac{Q_2}{w} = \kappa\right)$ , and setting  $Q_1 = 0$  ( $Q_1$  determines rate but not direction of selection), we solve for  $w$  in the substituted formula; we solve for  $\rho$  since  $\rho = w$  in this case. The actual calculations await the implementation of the SAC-1 polynomial system at Cornell.

A word on logic: Boolean expressions using  $\neg$ ,  $\wedge$ ,  $\vee$ ,  $SD$ , and  $=$  for operators are a convenient way of keeping notation clean. There is no provision for them, however, in FOCAL in the PDP/8 computer (which costs the user nothing to run, and yields a generation every 2.3 minutes) or in SAC-1 polynomial system. There are logic systems, but it is hard to mix logical arithmetic with real number arithmetic unless the appropriate functions exist to map  $F$  and  $T$  into 0 and 1. The solution employed here is to substitute numeric functions for their logical counterparts; e.g.,  $X = Y$  is equivalent to  $1 - (X - Y)^2$  and  $I \text{ sd } J = SD \cdot (I - J)^2 + I \cdot J$  ( $X = 0,1$ ;  $Y = 0,1$ ;  $I = 0,1$ ;  $J = 0,1$ ). Thus expanding Equations [12] and [13] we get

$$\begin{aligned} 0 &\leq w < \infty \\ 0 &\leq Q_1 < \infty \\ 0 &\leq Q_2 < \infty \end{aligned} \quad [16]$$

$$\begin{aligned} N'_{\alpha,\beta,\gamma} = & \sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \sum_{B=0}^1 \left( w \cdot (SD \cdot (I - J)^2 + I \cdot J) + Q_1 + Q_2 \right. \\ & \cdot (1 - (SD \cdot (K - J)^2 + K \cdot J)) \cdot N_{I,J,K} \cdot |\beta - \phi| \\ & \cdot (1 - (\alpha - (I \cdot A + J(1 - A)))^2) \\ & \cdot (1 - (\gamma - (I \cdot B + J \cdot (1 - B)))^2) / 8. \end{aligned}$$

$$\phi' = \frac{\sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \sum_{A=0}^1 \left( w \cdot (SD \cdot (I-J)^2 + I \cdot J) + Q_1 + Q_2 \cdot (1 - (SD \cdot (K-J)^2 + K \cdot J)) \right) \cdot N_{I,J,K} \cdot (1 - (I \cdot A + J \cdot (I-A)))}{\sum_{I=0}^1 \sum_{J=0}^1 \sum_{K=0}^1 \left( w \cdot (SD \cdot (I-J)^2 + I \cdot J) + Q_1 + Q_2 \cdot (1 - (SD \cdot (K-J)^2 + K \cdot J)) \right) \cdot N_{I,J,K}}$$

[17]

REFERENCES

1. BERRY, P. [1969] Form Number GH 20-0689-1 APL/360 PRIMER, Program Nos. 5734-X41, 5736-XM1. Published by IBM.
2. CROZIER, R. H. [1970] Coefficients of relationship and the identity of genes by descent in the Hymenoptera. An. Nat. Vol. 104, pp. 216-217.
3. FALCONER, P. S. [1961] Introduction to quantitative genetics. Oliver and Boyd.
4. HAMILTON, W. D. [1964] The geometrical evolution of social behavior, I and II. Jour. Theoret. Biol. Vol. 7, pp. 1-17, 18-52.
5. IVERSON, K. E. [1962] A Programming Language, John Wiley and Sons.
6. KEMPTHORNE, O. [1957] An Introduction to Genetic Statistics, John Wiley and Sons.
7. LI, C. C. [1955] Population Genetics, University of Chicago Press. See pp. 18-19 for discussion of independent genes.
8. MALICOTT See Maynard Smith, 1968.
9. MAYNARD SMITH, J. [1964] Group selection and kin selection. Nature, Vol. 201, pp. 1145-1147.
10. MAYNARD SMITH, J. [1968] Mathematical Ideas in Biology, Cambridge University Press.
11. TABER, S. [1955] Sperm distribution in spermatheca of mated queen honeybees. Apis Mellifera L. Hymenoptera. Jour. Econom. Ento. Vol. 45, No. 5, p. 522.
12. WELLS, H. G., WELLS, G. P., and HUXLEY, J. [1931] The Science of Life.
13. WEST, M. J. [1967] Foundress associations in Polistine wasps. Dominance hierarchies and the evolution of social behavior. Science, Vol. 157, pp. 1584-1585.
14. WRIGHT [1959] Encyclopedia Britannica. Look up Genetics of Populations.